Frequently Asked Questions for Users of the BioSense Application

BioSense Application Overview

1. **How does the BioSense application fit in with the Public Health Information Network (PHIN)?**
   Through defined data and vocabulary standards and strong collaborative relationships, PHIN will enable consistent exchange of response, health, and disease tracking data among public health partners. Ensuring the security of this information is critical as is the ability of the network to work reliably in times of national crisis. PHIN is composed of five key components: early event detection, outbreak management, connecting laboratory systems, countermeasure and response administration, and partner communications and alerting. BioSense is the early event detection component within the PHIN framework.

2. **How will the Outbreak Management System (OMS) and other PHIN systems integrate with the BioSense application?**
   Part of the BioSense Initiative is to advance early detection data standards that will work with outbreak management data standards under the broad umbrella of the Public Health Information Network (PHIN). It is critical for systems like BioSense to have the ability to share early event detection data with other systems (like the Outbreak Management System and other appropriate PHIN systems) to ensure that early detection leads to outbreak management and appropriate countermeasure administration. We plan to have data from BioSense be exchangeable with OMS.

3. **There are several early event detection systems available. Which one should I use?**
   We encourage you to try a variety of systems and evaluate which system or systems best meet your particular early detection and surveillance needs. We advocate the adoption of the PHIN standards by all early event detection and surveillance systems. This will help to foster system interoperability and data exchange among systems. There are a number of early detection activities which may be supportive of or supported by BioSense in the near future. These systems may provide opportunities for enhancing the data needs of early detection as well as establishing partnerships in pursuing methods for supplying useful public health data. Substantial work still needs to be done in identifying and optimizing the best algorithms for early detection. At this point, it is critical for all systems to use appropriate data and technology standards so data can be shared with all participating public health jurisdictions.

4. **Could the BioSense application be implemented in my local region?**
   BioSense is a web-based application that can be made available to any state or local public health entity. The system is national in scope, but is capable of presenting regional views so a local jurisdiction can easily access specific data and analyses for the particular region they are concerned with. BioSense currently includes over 30 major metropolitan areas, and users may request the creation of a “customized” MRA to meet surveillance needs. A region needs to be represented by an adequate amount of data in order for the analysis output to be useful. Acquisition of new data sources continues to be a top priority in order to better serve communities nationwide.

5. **How does a state or local department partner in the development of the BioSense application?**
   There are opportunities for states and locals to partner in the development of BioSense by providing feedback regarding the use of the system for monitoring, useful system enhancements, and additional analytic techniques. State and local health departments interested in partnering with BioSense should contact BioSenseHelp@cdc.gov.

6. **What are the 11 syndrome categories and how are the syndromes defined?**
   The data in BioSense are categorized as pre-diagnostic or syndromic. A multi-agency working group was established to identify and define eleven candidate syndrome groups. Definitions for each syndrome group were created by consensus; after an exhaustive search through all possible candidates, selected ICD-9-CM codes were categorized in one or more syndrome groups. A complete syndrome list and more detailed information on the syndrome categories are available at http://www.bt.cdc.gov/surveillance/syndromedef/index.asp.

7. **Does the BioSense application include national data only, or are there plans for global surveillance or for surveillance in U.S. territories such as Puerto Rico?**
   The first priority of BioSense is to cover states and major metropolitan areas in the United States. BioSense currently includes all U.S states and territories including Puerto Rico, Guam, Virgin Islands, and Marshall Islands, as well as over 30 major metropolitan areas. It is our intent to phase expansion in access for territories as appropriate data become available.

8. **What data formats can BioSense receive? How are the data feeds secured?**
   BioSense uses PHIN standards (messages, terminology, transport, and security) for data transmission. Data feeds into BioSense are securely transmitted daily utilizing the BioRetreiver with Public Health Information Network Messaging System (PHIN MS) and File Transfer Protocol (FTP).

9. **Does BioSense receive test results data from the BioWatch environmental samplers?**
   Yes, BioWatch results are available within BioSense. As part of the BioWatch initiative, lab results from the BioWatch collectors are being generated in Laboratory Response Network (LRN) labs. CDC developed an application for those labs that generates an HL7 [Health Level 7] standard lab result message for Category A agents. Hence, BioSense has the capability to deliver electronic lab results from LRN labs to public health jurisdiction from which that lab result was generated.

10. **Is the BioSense application sensitive enough to discern a small event (e.g., the 2001 anthrax attacks) from other daily data fluctuations?**
    Different types of outbreaks can be detected in different ways. For many outbreaks, an astute clinician may be the best detector. In BioSense, there are analytical indicators for significant data fluctuations to aid in the detection of medium to large potential outbreak situations. The BioSense application can also assist with the detection of smaller events

in the Sentinel Infection Alerts section. This alert posts a "warning box" on the BioSense home page of any state or MRA with a recent record of a provisional disease diagnosis that could potentially indicate a serious disease outbreak or bioterrorist event. Regardless of the initial method of detection, these BioSense approaches have value in subsequent detection, quantification, and localization of an event.

11. What software is used to analyze data?

To prepare the data for presentation, they are converted into cubes within SAS™ for quicker processing. BioSense utilizes the analytical components of the SAS™ software along with the CuSum and SMART algorithm techniques which are programmed specifically for the analysis of BioSense data.

12. How co-linear are the CuSUM and SMART score algorithms?

The CUSUM approach and SMART score approaches differ. The CUSUM implicitly assumes a moving constant baseline, while the SMART regression explicitly assumes the regression model over the cumulative historical period for its baseline. Suppose that we were to see observed counts that were identical from SMART model predictions and suppose it is the time of year when we have the onset of flu season and counts are increasing with time. Then the SMART scores would not produce an alert because the increase in counts is normal according to its modeling [the inclusion of last year's flu season into the modeling already internalized the previous unusual activities and may not see new high level of counts as unusual]. But the CUSUM might alert because it sees the deviation from a constant baseline. CUSUM looks at State/MRA level as unit of analysis while SMART looks at ZIP level as unit of analysis. So, a process of aggregation from ZIP to higher level is needed for SMART to have a direct comparison with CUSUM. However, this mapping is under scrutiny.

13. Can you define a "SMART" score?

The 'SMART' Scores in BioSense are calculated using a model-based approach that utilizes historical data for the zip code or area observed. A prediction is calculated daily using a Generalized Linear Mixed Model specifically for each area and this prediction is compared to the actual data received for the day, allowing a "SMART" score to be determined. The algorithm for this "SMART" score was provided by Ken Kleinman, Harvard Medical School, Harvard Pilgrim Health Care and Harvard Vanguard Medical Associates. The value of the 'SMART' Score is based on a transformation of a calculated p-value for comparing a prediction with the actual observed data count.

14. **How much baseline data are required for SMART scores?**

Baselines can be calculated with a few months of data to take 'day-of-the-week' into consideration, but require a couple of years to minimally account for 'seasonability' or 'month' effects. The SMART scores are currently implemented as an intermediate surveillance algorithm using several months of data (BioSense is a relatively young organization). Some data sources have several years of historical data included for our use. When this extensive data is loaded into the data mart, SMART scores may be generated that use additional terms in their models to account for 'seasonability'.

BioSense Application Access, Training, and Support

1. **Who can access the BioSense application, and how is access obtained?**
   State and local public health officials may access BioSense. Each state and metropolitan area (MRA) has a BioSense administrator identified by public health officials in that jurisdiction. The administrator may approve other public health users for access to data for the appropriate jurisdiction within BioSense. If you need assistance identifying and contacting your BioSense administrator, please contact the BioSense help desk at BioSenseHelp@cdc.gov for further assistance. Upon approval from the BioSense administrator, BioSense access can be obtained by applying and being approved for a digital certificate through the CDC Secure Data Network (http://sdn.cdc.gov).

2. **How does the BioSense application define regions and how do we out find if data is being collected for our region?**
   Theoretically, a BioSense "region" (state or MRA) can be defined for any grouping of zip codes that a public health department has appropriate authorization to view. Realistically, such a region would need to be represented by an adequate amount of data in order for BioSense output to be meaningful. Public health users at the local level have the option of requesting a "customized" MRA region to support their surveillance activities. To learn more about BioSense data coverage in your region, please email BioSenseHelp@cdc.gov.

3. **How does the BioSense application structure access to state and MRA level data?**
   The BioSense Initiative has worked with state and local health departments to identify appropriate data sharing guidelines. These public health partners were most comfortable with data being shared in association with specific jurisdictional needs. The default policy for BioSense will have MRA data being shared with all local jurisdictions that participate in that metropolitan geographic area. State health departments will have access to data related to their state, including any MRAs that fall completely within state borders. As part of the BioSense system, we have also agreed to facilitate modifications to these jurisdictional policies on the basis of overlapping and contiguous metropolitan areas. Other agencies needing access to the BioSense system will receive authorization from the appropriate jurisdictional BioSense Administrator(s).

4. **Can I access the BioSense application from any computer?**
   BioSense is a web-based system and can be accessed via the internet. The only requirement is that your digital certificate must be installed on the computer you use to access the system. You may export/import your certificate from your desktop to your laptop and vice versa if you would like to use more than one computer to access BioSense.

5. **Is documentation available to assist with using the BioSense application?**
   BioSense has a robust and frequently updated online help system, which includes documentation on the various data sources, analytic methods, and visualization techniques employed in the application.

BioIntelligence Center

1. **What is the BioIntelligence Center (BIC) at CDC?**
   The BioIntelligence Center maintains the daily operations of BioSense, including:

   - Ongoing monitoring and evaluation of BioSense data
   - Data anomaly investigation and tracking
   - Providing monitoring support to state and local health officials
   - Coordinating state and local public health training, support, and communications
   - Application troubleshooting
   - Generating ideas for system enhancements

   If you would like BIC monitoring support, please email BioSenseHelp@cdc.gov.

2. **Is CDC providing resources for monitoring the BioSense application, or is CDC relying on public health jurisdictions to do the monitoring using this tool?**
   The CDC staffs the BioIntelligence Center (BIC) to analyze data and assist local and state health departments with their analyses. In addition to these activities, the staff operates an "analytical help desk" to assist state and local users in the identification and investigation of data anomalies found. Because most public health surveillance and investigations occur at the state and local level, BioSense will be a useful adjunct to these activities. Recognizing resource limitations in public health nationally, a major thrust of BioSense is to minimize the state and local burden of data monitoring, analysis and investigation.

Research

1. **Is monitoring the BioSense application a research activity, and do these activities require Institutional Review Board (IRB) review?**
   The reporting of data in BioSense is not a research activity; it is public health practice that is consistent with the rules and regulations of HIPAA. The data sources provide data either with a data use agreement under the limited data set rules of HIPAA or through authorizations for disclosure pursuant to the public health provisions of the Privacy Rule.
2. **Are there published studies that demonstrate a correlation between perturbations in surveillance for broad syndrome categories and true outbreaks?**
   The BioSense initiative and application are intended to advance near real-time surveillance capabilities and assist public health departments in monitoring the health status of their jurisdictions. An extensive research listing regarding the use of early event detection systems is available at http://www.cdc.gov/epo/dphsi/syndromic.htm.